Mendelian randomization estimates of alanine aminotransferase with cardiovascular disease: Guangzhou Biobank Cohort Study

Xu, Lin; Thomas, G Neil; Cheng, Kar

DOI:
10.1093/hmg/ddw396

Document Version
Peer reviewed version

Citation for published version (Harvard):
Xu, L, Thomas, GN & Cheng, K 2017, 'Mendelian randomization estimates of alanine aminotransferase with cardiovascular disease: Guangzhou Biobank Cohort Study', *Human Molecular Genetics*, vol. 26, no. 2, pp. 430-437. https://doi.org/10.1093/hmg/ddw396

Mendelian randomization estimates of alanine aminotransferase with cardiovascular disease: Guangzhou Biobank Cohort Study

Lin Xu,[1] Chao Qiang Jiang,[2] Tai Hing Lam,[1,2*] Wei Sen Zhang,[2] Feng Zhu,[2] Ya Li Jin,[2] G Neil Thomas,[3] Kar Keung Cheng,[3] C Mary Schooling[1,4]

[1]School of Public Health, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong SAR, China;

[2]Guangzhou 12th Hospital, Guangzhou, China;

[3]Institute of Applied Health Research, University of Birmingham, Birmingham, UK

[4]CUNY School of Public Health at Hunter College, 2180 Third Avenue, New York, NY 10035, USA

[*] Corresponding author

Professor Tai Hing Lam

School of Public Health, The University of Hong Kong

Hong Kong, China

Tel: (852) 3917 9287; **Fax:** (852) 2855 9528

Email: hrmrlth@hku.hk

Abstract

Observational studies of the association of alanine aminotransferase (ALT) levels with ischemic heart disease (IHD) and cardiovascular disease (CVD) risk factors are inconsistent, probably because of confounding and reverse causality. Mendelian randomization (MR) provides less confounded results. We used MR analysis to assess the associations of ALT (U/L) with IHD, diabetes and other CVD risk factors. We used instrumental variable analysis based on two single nucleotide polymorphism (SNPs) *HSD17B13/MAPK10* (*rs6834314*) and *PNPLA3/SAMM50* (*rs738409*) to assess the associations of ALT (U/L) with IHD, diabetes and other CVD risk factors in the Guangzhou Biobank Cohort Study (GBCS). Observationally in 19,925 participants ALT levels were strongly positively associated with self-reported IHD, systolic and diastolic blood pressure, low-density lipoprotein- and total cholesterol, triglycerides, fasting glucose, body mass index, waist circumference, heart rate (HR) and diabetes, but were not associated with uncorrected QT interval, HR-corrected QT interval or high-density lipoprotein-cholesterol. In the MR study, using a credible genetic instrument (*F*-statistic=23) for ALT, ALT levels were negatively associated with IHD (odds ratio (OR) 0.92, 95% confidence interval (CI) 0.87 to 0.97) and triglycerides (β -0.08, 95% CI -0.13 to -0.03), but were not associated with other CVD risk factors. Our results using Mendelian randomization suggest that ALT reduces the risk of IHD, probably through reducing triglyceride levels. The underlying mechanisms deserve further investigation.

Introduction

Serum alanine aminotransferase (ALT) is a non-specific marker for liver fat. High ALT levels indicate the release of aminotransferase from cytoplasm to blood stream probably due to damaged liver (1, 2) and are associated with a higher risk of diabetes.(3) Prospective studies of the association of ALT with cardiovascular disease (CVD) showed inconsistent results.(4, 5) Meta-analyses of prospective cohort studies report an inverse association of ALT levels with CVD mortality, despite great heterogeneity (I^2 range 79-82).(4, 5) The prospective associations also appeared to vary by region and age, with estimates suggestive of positive associations in Asians and younger people but possible negative associations in people from Europe or North America and older people.(4, 5) Higher ALT levels are associated with incident type 2 diabetes.(3, 6, 7) Given diabetes is a known risk factor for CVD, it is unclear whether these divergent associations are the result of confounding by factors related to aging (such as frailty and reduced skeletal muscle mass),(8-10) poor nutrition and other biomarkers of hepatocyte function, which are often difficult to measure and control for in traditional observational studies,(5) or another example of a factor, such as statins(11) and familial hypercholesterolemia,(12, 13) with opposite effects on IHD and diabetes.

Increasing interest in the potential value of ALT in CVD prevention means MR study is a vital first step before investigating in depth the underlying mechanisms in an experimental approach. Moreover, MR study tests a causal pathway, whilst trials of pharmacologic agents or interventions for improving liver function that influence ALT but may also influence other

several hepatic and lipid factors,(14, 15) making it difficult to distinguish the effects attributable to ALT from other off-target effects. No intervention that specifically modifies ALT is known. Examining the effect of ALT using Mendelian randomization (MR) is the most appropriate, cost-effective and timely approach to assess the effects of ALT on IHD and its risk factor. Here, we conducted the first MR study of the effect of ALT on IHD and its risk factors based on individual-level data from the Guangzhou Biobank Cohort Study.

Results

Of the four ALT-related SNPs (rs6834314, rs2954021, rs10883437, rs738409) from a recent GWAS(16) tested in 10,623 older Chinese people in GBCS with mean age of 61.5 (standard deviation (SD) 6.9) years, two SNPs (rs2954021 and rs10883437) were excluded. Rs2954021 had a total missing rate >90% and minor allele frequency <0.05, and deviated from Hardy-Weinberg Equilibrium. Rs10883437 was not associated with ALT levels (p-value 0.73). Two SNPs, *HSD17B13/MAPK10* (*rs6834314*) and *PNPLA3/SAMM50* (*rs738409*), were used for constructing the allele score. The effects of these two variants on plasma ALT in our sample were smaller than those reported in the GWAS (Appendix Table 1). These two genetic variants were defined as being independent of each other on the basis of low correlation ($R^2 < 0.1$) in HapMap22 or the 1000 genome project data. A weighted genetic score was created based on the equation of "0.6×rs6834314+0.6×rs738409+24", with the genotypes being coded into 0, 1 and 2, and the number of alleles was used as continuous variables. The *F*-statistic from the first stage of the IV analysis was 23, suggesting that weak-instrument bias was unlikely.

Table 1 shows that ALT levels were negatively associated with age, being a woman and alcohol use, and positively associated with education, waist circumference and BMI. No clear association with smoking or physical activity was found (Table 1). As expected, most of the CVD risk factors were correlated with each other (Appendix table 2), while neither rs6834314 nor rs738409 was associated with socioeconomic position or lifestyle, including age, sex, education, smoking status, use of alcohol and physical activity, suggesting these SNPs were not affected by confounding (Table 2).(17)

In 19,925 GBCS participants, adjusted for age, sex, education, smoking status, alcohol use, physical activity, BMI and waist circumference (except for the results for BMI and waist circumference, respectively), ALT was positively associated with systolic and diastolic blood pressure, total and LDL-cholesterol, triglycerides, fasting glucose, BMI, waist circumference, heart rate, and diabetes (all P values <0.001), but not with HDL-cholesterol levels, corrected QT interval, uncorrected QT interval or self-reported IHD (Table 3).

Table 4 shows that, in the MR study, ALT was negatively associated with triglyceride levels (β -0.08, 95% confidence interval (CI) -0.13 to -0.03) and consistently negatively associated with IHD (OR 0.92, 95% CI 0.87 to 0.97). No association with other CVD risk factors, including systolic and diastolic blood pressure, HDL-, LDL- and total cholesterol, fasting glucose, BMI, waist circumference, HR-corrected and uncorrected QT interval, HR, or type 2 diabetes were evident (p-value from 0.21 to 0.98). MR estimates for blood pressure and

cholesterol were of similar magnitude and direction as the observational estimates, whilst the MR estimates for BMI, WHR and QT interval were more different from the observational estimates. Sensitivity analysis additionally including rs10883437 in the genetic instrument showed similar results (Appendix table 3).

Discussion

Using a Mendelian randomization study in an under-studied population, ALT levels were negatively associated with IHD, probably via lowering triglyceride levels. The results for IHD were consistent with the emerging body of evidence from large scale prospective cohort studies,(9, 18) and meta-analysis of older adult samples.(4) Using Mendelian randomization ALT had few other associations with other CVD risk factors. In contrast with the MR results observationally ALT was positively associated with self-reported IHD and CVD risk factors most likely to residual confounding.

The positive association of ALT with type 2 diabetes was consistent in observational analysis with and without adjustment. The MR association of ALT with type 2 diabetes was almost null, which may be because type 2 diabetes in the GBCS was defined according to fasting glucose levels and self-reports only. The lack of other factors that can contribute to the clinical diagnosis, i.e., taking into account diabetes related symptoms, blood glucose after an 2-hour oral glucose tolerance test, HbA1c, and/or a repeat testing on a second occasion, may lead to misclassification (under-diagnosis) of the type 2 diabetes and dilute the gene-disease association, which will attenuate the MR estimate toward the null. However,

the definition of type 2 diabetes in the GBCS is usually used in large epidemiologic studies. However, the MR estimate for fasting glucose was directionally and quantitatively similar to the observational estimate, but had a confidence interval including the null because of the larger sample size needed for MR studies. As such an adverse effect of ALT on fasting glucose cannot be ruled out.

An MR study approach provides relatively less confounded estimates of causal effects, given unavoidable confounding factors from other hepatic factors in traditional observational studies. Potential limitations in MR studies, such as confounding from linkage disequilibrium and population stratification, pleiotropy and canalization should not be major concerns in our study. Firstly, the existence of pleiotropy, where a genetic instrumental variable has an effect on an outcome (CVD risk factors, IHD or diabetes) independent of its effect on the exposure (ALT) would have implications for the assumptions made in the MR analysis. However, we assessed potential pleiotropy effects using the Ensembl gene annotation system (19) and found no evidence for the existence of other phenotypes for these genetic instruments. Similarly, if a genetic variant in the score was in linkage disequilibrium with another genetic variant that influences the outcome through a pathway that is unrelated to the exposure, this could also bias the causal estimate. However, no linkage disequilibrium between the ALT-related SNPs in the GWAS(16) or from the SNP Annotation and Proxy Search system (20) was identified. In addition, the consistency of IV estimates obtained using two allele scores in this study suggests that pleiotropy is unlikely. Secondly, population stratification, which refers to the existence of the differences in allele

frequencies and disease prevalence in different ethnic groups, is also unlikely because we included permanent residents of Guangzhou who were homogeneous Southern Chinese. Thirdly, we used the most functionally relevant SNPS identified from a large GWAS as the genetic instruments, and no other phenotypic traits or biological effects related to IHD or type 2 diabetes for these 2 SNPs have been reported. Fourthly, we cannot completely rule out the effect of canalization in MR studies. Sixthly, it would be very useful to examine the temporal stability in the estimates of relationship between SNPs and ALT and for the MR analyse. However, we did not have sufficient number of participants with repeated measurements of ALT to conduct such analyses. Further studies with repeated measurements of ALT are warranted to examine the stability of the genetic instruments. Additionally, MR studies require large sample sizes. Meta-analysis of MR studies may provide more precise estimates of the effect size. But no other MR studies on ALT were found. Further two-sample MR studies using consortium-based GWAS (21) or one-sample MR in large cohorts (i.e. UK biobank) are warranted to replicate our results. Seventhly, we used 2SLS for the MR analysis, in which the output from the first-stage regression was fed into the second-stage regression with no acknowledgement of uncertainty.(22) Finally, as some CVD risk factors (i.e. adiposity (23) or lipids (24)) were also associated with ALT, further MR studies using functionally related loci are helpful in addressing the potential issue of reverse causation (i.e. elevated CVD risk factors increase ALT level).

Although the inverse association of ALT with IHD is convincing, the mechanistic pathways are not fully understood. Our results suggest one possible pathway is through the effect on

triglycerides. Moreover, it has been suggested that low ALT levels may reflect impaired synthetic capacity of the liver.(25) As hepatocytes play important roles in detoxification and lipid metabolism, a decrease in functional hepatocytes may increase the susceptibility to toxins and metabolic disorders, which leads to higher risks of IHD.(9) Low ALT levels could also be a proxy for liver aging(26)and aging-related reduced skeletal muscle mass,(9) because a small amount of ALT is derived from skeletal muscle.(27) ALT levels might also be regulated by androgens, which may have deleterious effects on cardiovascular disease(28) and lipids,(29) but are inversely associated with glucose probably via increasing skeletal muscle mass.(30) An animal study showed ALT gene expression was affected by castration or androgen administration in non-hepatic tissues.(31) Further studies to clarify whether androgen plays a role in the ALT and IHD/diabetes associations are warranted.(10) Finally, the discrepancy between the observational and MR findings are illuminating, because lack of replication of observed relations suggests strong confounding by causal factors that might be of interest in their own right. For example observationally obesity strongly predicted higher ALT, but in MR the relation was in an opposite direction, indicating causal factors that increase both adiposity and ALT.

Our results using Mendelian randomization suggest that ALT reduces the risk of IHD, whilst an adverse effect on fasting glucose cannot be excluded. The implications of these findings for the primary prevention of CVD, such as lifestyle interventions or use of pharmacological agents, warrant further investigation.

Materials and Methods

Participants

The Guangzhou Biobank Cohort Study (GBCS) is a 3-way collaboration among Guangzhou 12th Hospital and the Universities of Hong Kong and Birmingham, UK, which has been described in detail elsewhere,(32) and several papers on MR have been published.(29, 33, 34) The GBCS baseline examination was conducted in three phases from 2003 to 2008, in the Guangzhou 12th Hospital, included an interview on lifestyle, family and personal medical history and assessment of anthropometric and clinical factors. Information on socioeconomic position and lifestyle including age, sex, education, smoking and alcohol use was collected by a computer-assisted standardized questionnaire administered by trained interviewers. Physical activity was assessed using a validated Chinese version of the International Physical Activity Questionnaire.(35) Anthropometric measurements were performed by trained nurses using standard protocols. Participants wore light clothing and no shoes. Body weight was measured to the nearest 0.1 kilogram using a platform scale (RGZ-120-RT, China). Waist circumference was measured horizontally around the smallest circumference between ribs and iliac crest, or at the navel, if there was no natural waistline. Body mass index (BMI) was calculated using measured weight and height as weight in kilograms divided by height in meters squared. Plasma glucose, lipids and liver enzymes were measured by Shimadzu CL-8000 Clinical Chemistry Analyzer (Shimadzu, Kyoto, Japan). All measurements of biochemical parameters from 2003 to 2008 were conducted in the same laboratory of the Guangzhou Number 12 Hospital using the same methods. A standard electrocardiogram (ECG) was also performed The details of the ECG measurement have been

reported elsewhere.(33, 34)The Guangzhou Medical Ethics Committee of the Chinese Medical Association approved the study and all participants gave written, informed consent before participation.

DNA extraction and single nucleotide polymorphism (SNP) analysis

DNA was extracted at Guangzhou Number 12Hospital either from fresh blood using a standard phenol-chloroform extraction procedure and stored at -80°C or from blood or buffy coat previously stored at -80°C using a standard magnetic bead extraction procedure. Genotyping was performed using the MassARRAY system (Sequenom, San Diego, CA, USA) at a commercial company (Beijing Genomics Institute, Shenzhen, China).

Exposure and outcomes

Data for all exposure and outcome variables were collected at the baseline examination. The exposure was ALT (U/L). The outcomes included self-reported IHD, type 2 diabetes mellitus (T2DM), and cardiovascular disease risk factors including total, HDL- and LDL-cholesterol, systolic and diastolic blood pressure, fasting glucose, BMI, waist circumference, QT interval, heart rate (HR) corrected QT intervals using the Framingham formula (calculated as QT+154×(1-60/HR))(36) and HR. T2DM was defined as self-reported physician diagnosis diabetes, use of hypoglycemia medication or insulin, and/or fasting glucose ≥7.0mmol/l.(37)

Instrumental variable (IV) for ALT

Four SNPs (*rs6834314, rs2954021, rs10883437, rs738409*)associated with plasma ALT

concentrations at P < 1×10^-8 in a recent GWAS(16) were genotyped in the GBCS. The SNPs significantly associated with ALT in our sample were used as an IV for ALT. Allele scores were derived using the dose of the effect allele at each SNP which was first weighted by the effect size of the variant and then summed:

$$\text{Weighted ALT score} = w_1 \times SNP_1 + w_2 \times SNP_2 + \cdots w_n \times SNP_n$$

Where w is the weight (i.e. the beta-coefficient of association of the SNP with ALT) and SNP is the dosage of ALT-raising alleles at that locus (i.e. 0, 1 or 2 ALT raising alleles).

Potential Confounders

Potential confounders, selected as common causes of ALT and the outcomes, included age, sex, education, smoking, alcohol use, physical activity, BMI and waist circumference were used in the observational analysis.

Statistical analysis

We tested for Hardy-Weinberg equilibrium at the SNP locus on a contingency table of observed-versus-predicted frequencies with an exact test. We used analysis of variance (ANOVA) and linear regression to assess the association of potential confounders with ALT. Linear regression with bootstrapping internal validation was used to select the SNPs which best predicted ALT. An allele score was created by summing the number of risk alleles with weighting.(38) We used 2 stage least squares (2SLS) to estimate the possible causal effect of ALT on each outcomes, i.e., the change in health outcomes per U/L increase in ALT. To avoid weak IV bias, we checked the *F*-statistic from the first stage was greater than 10,

which indicates that the IV is unlikely to be weak.(39) We also used Durbin-Wu-Hausman endogeneity test to test for endogeneity in a regression estimated with IV. The null hypothesis is that an ordinary least squares (OLS) estimator of the same equation would yield consistent estimates. We also conducted sensitivity analysis using all variants available for the genetic instrument. All statistical analysis was performed using STATA 13.1 (Stata Corp LP, College Station, TX, USA).

Acknowledgments: The Guangzhou Biobank Cohort Study was funded by The University of Hong Kong Foundation for Educational Development and Research (SN/1f/HKUF-DC;C20400.28505200), the Health Medical Research Fund (Grant number: HMRF/12133631) in Hong Kong; the Guangzhou Public Health Bureau (201102A211004011) and the Guangzhou Science and Technology Bureau, Guangzhou, China (2002Z2-E2051; 2012J5100041; 2013J4100031); and the University of Birmingham, UK.

Conflict of Interest Statement: none

References

1 Pratt, D.S. and Kaplan, M.M. (2000) Evaluation of abnormal liver-enzyme results in asymptomatic patients. *The New England journal of medicine*, **342**, 1266-1271.

2 Sherman, K.E. (1991) Alanine aminotransferase in clinical practice. A review. *Arch Intern Med*, **151**, 260-265.

3 Kunutsor, S.K., Apekey, T.A. and Walley, J. (2013) Liver aminotransferases and risk of

incident type 2 diabetes: a systematic review and meta-analysis. *American journal of epidemiology*, **178**, 159-171.

4 Liu, Z., Ning, H., Que, S., Wang, L., Qin, X. and Peng, T. (2014) Complex association between alanine aminotransferase activity and mortality in general population: a systematic review and meta-analysis of prospective studies. *PLoS One*, **9**, e91410.

5 Kunutsor, S.K., Apekey, T.A. and Khan, H. (2014) Liver enzymes and risk of cardiovascular disease in the general population: a meta-analysis of prospective cohort studies. *Atherosclerosis*, **236**, 7-17.

6 Xu, L., Jiang, C.Q., Schooling, C.M., Zhang, W.S., Cheng, K.K. and Lam, T.H. (2015) Liver enzymes and incident diabetes in China: a prospective analysis of 10 764 participants in the Guangzhou Biobank Cohort Study. *J Epidemiol Community Health*, **69**, 1040-1044.

7 Nguyen, Q.M., Srinivasan, S.R., Xu, J.H., Chen, W., Hassig, S., Rice, J. and Berenson, G.S. (2011) Elevated liver function enzymes are related to the development of prediabetes and type 2 diabetes in younger adults: the Bogalusa Heart Study. *Diabetes Care*, **34**, 2603-2607.

8 Afilalo, J., Karunananthan, S., Eisenberg, M.J., Alexander, K.P. and Bergman, H. (2009) Role of frailty in patients with cardiovascular disease. *Am J Cardiol*, **103**, 1616-1621.

9 Ford, I., Mooijaart, S.P., Lloyd, S., Murray, H.M., Westendorp, R.G., de Craen, A.J., Packard, C.J., Buckley, B., Barlow, C. and Preiss, D. (2011) The inverse relationship between alanine aminotransferase in the normal range and adverse cardiovascular and non-cardiovascular outcomes. *International journal of epidemiology*, **40**, 1530-1538.

10 Schooling, C.M., Kelvin, E.A. and Jones, H.E. (2012) Alanine transaminase has

opposite associations with death from diabetes and ischemic heart disease in NHANES III. *Ann Epidemiol*, **22**, 789-798.

11 Sattar, N., Preiss, D., Murray, H.M., Welsh, P., Buckley, B.M., de Craen, A.J., Seshasai, S.R., McMurray, J.J., Freeman, D.J., Jukema, J.W. *et al.* (2010) Statins and risk of incident diabetes: a collaborative meta-analysis of randomised statin trials. *Lancet*, **375**, 735-742.

12 Vuorio, A., Strandberg, T.E., Schneider, W.J. and Kovanen, P.T. (2016) Statins and new-onset diabetes mellitus – a risk lacking in familial hypercholesterolaemia. *Journal of Internal Medicine*, **279**, 358-361.

13 Besseling, J., Kastelein, J.J., Defesche, J.C., Hutten, B.A. and Hovingh, G.K. (2015) Association between familial hypercholesterolemia and prevalence of type 2 diabetes mellitus. *JAMA*, **313**, 1029-1036.

14 Musso, G., Gambino, R., Cassader, M. and Pagano, G. (2010) A meta-analysis of randomized trials for the treatment of nonalcoholic fatty liver disease. *Hepatology*, **52**, 79-104.

15 McKenzie, J., Fisher, B.M., Jaap, A.J., Stanley, A., Paterson, K. and Sattar, N. (2006) Effects of HRT on liver enzyme levels in women with type 2 diabetes: a randomized placebo-controlled trial. *Clin Endocrinol (Oxf)*, **65**, 40-44.

16 Chambers, J.C., Zhang, W., Sehmi, J., Li, X., Wass, M.N., Van der Harst, P., Holm, H., Sanna, S., Kavousi, M. and Baumeister, S.E. (2011) Genome-wide association study identifies loci influencing concentrations of liver enzymes in plasma. *Nature genetics*, **43**, 1131-1138.

17 VanderWeele, T.J., Tchetgen Tchetgen, E.J., Cornelis, M. and Kraft, P. (2014)

Methodological challenges in mendelian randomization. *Epidemiology*, **25**, 427-435.

18 Harada, P.H., Cohen, D.E., Rose, L. and Ridker, P.M. (2014) Alanine Aminotransferase Levels and Incident Cardiovascular Events. *Circulation*, **130**, A12037-A12037.

19 Aken, B.L., Ayling, S., Barrell, D., Clarke, L., Curwen, V., Fairley, S., Fernandez Banet, J., Billis, K., García Girón, C., Hourlier, T. *et al.* (2016) The Ensembl gene annotation system. *Database*, **2016**.

20 Johnson, A.D., Handsaker, R.E., Pulit, S.L., Nizzari, M.M., O'Donnell, C.J. and de Bakker, P.I.W. (2008) SNAP: a web-based tool for identification and annotation of proxy SNPs using HapMap. *Bioinformatics*, **24**, 2938-2939.

21 Lawlor, D.A. (2016) Commentary: Two-sample Mendelian randomization: opportunities and challenges. *International journal of epidemiology*, **45**, 908-915.

22 Burgess, S. and Thompson, S.G. (2015) *Mendelian randomization: methods for using genetic variants in causal estimation*. CRC Press.

23 Popov, V.B., Thompson, C.C., Kumar, N., Ciarleglio, M.M., Deng, Y. and Laine, L. (2016) Effect of Intragastric Balloons on Liver Enzymes: A Systematic Review and Meta-Analysis. *Dig Dis Sci*, **61**, 2477-2487.

24 Rinella, M.E. (2015) Nonalcoholic fatty liver disease: A systematic review. *JAMA*, **313**, 2263-2273.

25 Sherman, K.E. (1991) Alanine aminotransferase in clinical practice: A review. *Archives of Internal Medicine*, **151**, 260-265.

26 Elinav, E., Ackerman, Z., Maaravi, Y., Ben-Dov, I.Z., Ein-Mor, E. and Stessman, J. (2006) Low alanine aminotransferase activity in older people is associated with greater

long-term mortality. *J Am Geriatr Soc*, **54**, 1719-1724.

27 Hong, H.C., Hwang, S.Y., Choi, H.Y., Yoo, H.J., Seo, J.A., Kim, S.G., Kim, N.H., Baik, S.H., Choi, D.S. and Choi, K.M. (2014) Relationship between sarcopenia and nonalcoholic fatty liver disease: the Korean Sarcopenic Obesity Study. *Hepatology*, **59**, 1772-1778.

28 Xu, L., Freeman, G., Cowling, B.J. and Schooling, C.M. (2013) Testosterone therapy and cardiovascular events among men: a systematic review and meta-analysis of placebo-controlled randomized trials. *BMC Med*, **11**, 108.

29 Zhao, J., Jiang, C., Lam, T.H., Liu, B., Cheng, K.K., Xu, L., Au Yeung, S.L., Zhang, W., Leung, G.M. and Schooling, C.M. (2014) Genetically predicted testosterone and cardiovascular risk factors in men: a Mendelian randomization analysis in the Guangzhou Biobank Cohort Study. *International journal of epidemiology*, **43**, 140-148.

30 Hou, W.W., Tse, M.A., Lam, T.H., Leung, G.M. and Schooling, C.M. (2015) Adolescent testosterone, muscle mass and glucose metabolism: evidence from the 'Children of 1997' birth cohort in Hong Kong. *Diabet Med*, **32**, 505-512.

31 Coss, C.C., Bauler, M., Narayanan, R., Miller, D.D. and Dalton, J.T. (2012) Alanine aminotransferase regulation by androgens in non-hepatic tissues. *Pharm Res*, **29**, 1046-1056.

32 Jiang, C., Thomas, G.N., Lam, T.H., Schooling, C.M., Zhang, W., Lao, X., Adab, P., Liu, B., Leung, G.M. and Cheng, K.K. (2006) Cohort profile: The Guangzhou Biobank Cohort Study, a Guangzhou-Hong Kong-Birmingham collaboration. *Int J Epidemiol*, **35**, 844-852.

33 Au Yeung, S.L., Jiang, C., Long, M., Cheng, K.K., Liu, B., Zhang, W., Lam, T.H., Leung, G.M. and Schooling, C.M. (2015) Evaluation of Moderate Alcohol Use With QT Interval and Heart Rate Using Mendelian Randomization Analysis Among Older Southern

Chinese Men in the Guangzhou Biobank Cohort Study. *American journal of epidemiology*, **182**, 320-327.

34 Zhao, J., Jiang, C., Lam, T.H., Liu, B., Cheng, K.K., Xu, L., Long, M.J., Zhang, W., Leung, G.M. and Schooling, C.M. (2015) Genetically predicted testosterone and electrocardiographic QT interval duration in Chinese: a Mendelian randomization analysis in the Guangzhou Biobank Cohort Study. *International journal of epidemiology*, **44**, 613-620.

35 Xu, L., Jiang, C.Q., Lam, T.H., Zhang, W.S., Thomas, G.N. and Cheng, K.K. (2011) Dose-response relation between physical activity and cognitive function: guangzhou biobank cohort study. *Annals of epidemiology*, **21**, 857-863.

36 Sagie, A., Larson, M.G., Goldberg, R.J., Bengtson, J.R. and Levy, D. (1992) An improved method for adjusting the QT interval for heart rate (the Framingham Heart Study). *Am. J. Cardiol.*, **70**, 797-801.

37 Wilson, P.W., Meigs, J.B., Sullivan, L., Fox, C.S., Nathan, D.M. and D'Agostino, R.B., Sr. (2007) Prediction of incident diabetes mellitus in middle-aged adults: the Framingham Offspring Study. *Archives of internal medicine*, **167**, 1068-1074.

38 Burgess, S. and Thompson, S.G. (2013) Use of allele scores as instrumental variables for Mendelian randomization. *International journal of epidemiology*, **42**, 1134-1144.

39 Staiger, D. and Stock, J.H. (1997) Instrumental Variables Regression with Weak Instruments. *Econometrica*, **65**, 557-586.

Table 1. Baseline demographic characteristics in 19,925participants measured alanine aminotransferase (ALT) during 2003-2008 in Guangzhou Biobank Cohort Study

Characteristics	**Number**	**%**	**Alanine aminotransferase (U/L)**		
			Mean	**SD**	**P-value**[†]
Age group, years					
50-59	8,043	40.3	25.51	13.54	<0.001
60-69	8,394	42.2	25.24	12.27	
≥70	3,488	17.5	23.56	12.16	
Sex					
Women	14,471	72.6	24.39	12.49	<0.001
Men	5,454	27.4	26.81	13.41	
Education					
Primary school or less	8,758	44.0	24.67	12.44	<0.001
Secondary	9,413	47.2	25.19	13.00	
College or above	1,754	8.8	26.31	13.37	
Smoking status					
Never	16,070	80.7	24.97	12.73	<0.001
Former	1,827	9.2	26.57	12.91	
Current	2,028	10.2	24.38	13.09	
Alcohol use					
Never	13,701	68.8	25.04	12.63	0.03
Former	750	3.8	26.25	13.48	
Current	5,474	27.4	24.94	13.10	
Physical activity					
Inactive	916	4.6	24.59	13.84	<0.001
Moderately active	6,262	31.4	25.63	13.22	
Active	12,747	64.0	24.81	12.49	
			β	**95% CI**	**P-value**[‡]
Age, year	19,925	-	-0.08	-0.11 to -0.06	<0.001
Waist circumference, cm	19,925	-	0.32	0.31 to 0.34	<0.001
Body mass index, kg/m^2	19,925	-	0.85	0.80 to 0.90	<0.001

SD: standard deviation; CI: confidence interval

[†]: P values from analysis of variance;

[‡]: P values from linear regression

Table 2. Baseline demographic characteristics by ALT-related SNPs in 10,623 participants measured alanine aminotransferase (ALT) and ALT-related SNPs in Guangzhou Biobank Cohort Study

SNP	***HSD17B13/MAPK10* (*rs6834314*)**				***PNPLA3/SAMM50* (*rs738409*)**			
	GG(n=1,408)	**GA (n=4,918)**	**AA(n=4,297)**	**P-value**	**CC(n=4,959)**	**GC(n=4,508)**	**GG(n=1,156)**	**P-value**
ALT, U/L, mean (SD)†	24.5 (14.7)	25.0 (13.9)	25.7 (14.3)	0.009	24.9 (13.9)	25.2 (13.8)	26.5 (16.7)	0.002
Age, year, mean (SD)†	61.8 (6.9)	61.4 (6.9)	61.7 (7)	0.06	61.7 (7)	61.4 (6.9)	61.8 (7)	0.53
Sex, number (%)‡								
Women	1075 (76.4)	3737 (76)	3286 (76.5)	0.86	3797 (76.6)	3453 (76.6)	848 (73.4)	0.06
Men	333 (23.7)	1181 (24)	1011 (23.5)		1162 (23.4)	1055 (23.4)	308 (26.6)	
Education, number (%)‡								
Primary school or less	588 (41.8)	1945 (39.6)	1697 (39.5)	0.57	2024 (40.8)	1741 (38.6)	465 (40.2)	0.24
Secondary	705 (50.1)	2533 (51.5)	2212 (51.5)		2495 (50.3)	2359 (52.3)	596 (51.6)	
College or above	115 (8.2)	440 (9)	388 (9)		440 (8.9)	408 (9.1)	95 (8.2)	
Smoking status, number (%)‡								
Never	1148 (81.5)	4110 (83.6)	3605 (83.9)	0.11	4159 (83.9)	3767 (83.6)	937 (81.1)	0.17
Former	123 (8.7)	391 (8)	303 (7.1)		365 (7.4)	344 (7.6)	108 (9.3)	
Current	137 (9.7)	417 (8.5)	389 (9.1)		435 (8.8)	397 (8.8)	111 (9.6)	
Alcohol use, number (%)‡								
Never	952 (67.6)	3273 (66.6)	2938 (68.4)	0.30	3387 (68.3)	3002 (66.6)	774 (67)	0.37
Former	43 (3.1)	184 (3.7)	155 (3.6)		182 (3.7)	162 (3.6)	38 (3.3)	
Current	413 (29.3)	1461 (29.7)	1204 (28)		1390 (28)	1344 (29.8)	344 (29.8)	
Physical activity, number (%)‡								
Inactive	63 (4.5)	243 (4.9)	199 (4.6)	0.72	216 (4.4)	226 (5)	63 (5.5)	0.42
Moderately active	436 (31)	1459 (29.7)	1320 (30.7)		1501 (30.3)	1360 (30.2)	354 (30.6)	
Active	909 (64.6)	3216 (65.4)	2778 (64.7)		3242 (65.4)	2922 (64.8)	739 (63.9)	

Systolic blood pressure, mmHg, mean (SD)[†]	129.8 (21.8)	129.4 (21.3)	130.2 (21.6)	0.27	129.9 (21.6)	129.6 (21.6)	130 (20.9)	0.65
Diastolic blood pressure, mmHg, mean (SD)[†]	73.4 (11.3)	73.5 (11)	73.7 (10.9)	0.69	73.6 (10.9)	73.5 (11.1)	73.8 (10.8)	0.69
High density lipoprotein-cholesterol, mmol/l, mean (SD)[†]	1.7 (0.4)	1.7 (0.4)	1.7 (0.4)	0.46	1.7 (0.4)	1.7 (0.4)	1.7 (0.4)	0.72
Low density lipoprotein-cholesterol, mmol/l, mean (SD)[†]	3.3 (0.7)	3.3 (0.7)	3.2 (0.7)	0.31	3.2 (0.7)	3.3 (0.7)	3.3 (0.7)	0.06
Triglycerides, mmol/l, mean (SD)[†]	1.8 (1.4)	1.7 (1.1)	1.7 (1)	0.04	1.7 (1.2)	1.7 (1.1)	1.6 (1)	0.03
Total cholesterol, mmol/l, mean (SD)[†]	5.9 (1.2)	5.9 (1.1)	5.9 (1.1)	0.19	5.9 (1.1)	5.9 (1.1)	5.9 (1.1)	0.90
Fasting glucose, mmol/l, mean (SD)[†]	5.8 (1.4)	5.8 (1.4)	5.8 (1.6)	0.06	5.8 (1.5)	5.8 (1.5)	5.7 (1.4)	0.62
Body mass index, Kg/m^2, mean (SD)[†]	23.7 (3.3)	23.8 (3.3)	23.9 (3.3)	0.15	23.9 (3.3)	23.8 (3.3)	23.6 (3.1)	0.006
Waist circumference, cm, mean (SD)[†]	78.7 (8.8)	78.9 (8.9)	79.1 (9)	0.38	79.2 (9)	78.7 (9)	78.6 (8.7)	0.03
Corrected QT interval, milliseconds, mean (SD)[†]	412.1 (21.3)	412.3 (21.5)	412.3 (20.9)	0.95	411.7 (21.3)	413 (21.3)	411.9 (20.8)	0.02
Uncorrected QT interval, milliseconds, mean (SD)[†]	389.1 (28)	389.5 (28.9)	389.1 (28.2)	0.72	388.8 (28.7)	389.8 (28.4)	389.5 (27.8)	0.23
Heart rate, beats per minute, mean (SD)[†]	72 (10.6)	71.9 (10.6)	72.2 (10.8)	0.45	72 (10.8)	72.1 (10.6)	71.6 (10)	0.32
Ischemic heart disease, number (%)[‡]								
Yes	1352 (96)	4720 (96)	4144 (96.4)	0.49	4752 (95.8)	4352 (96.5)	1112 (96.2)	0.20
No	56 (4)	198 (4)	153 (3.6)		207 (4.2)	156 (3.5)	44 (3.8)	
Type 2 diabetes, number (%)[‡]								
Yes	1242 (88.2)	4355 (88.6)	3789 (88.2)	0.84	4370 (88.1)	3991 (88.5)	1025 (88.7)	0.78
No	166 (11.8)	563 (11.5)	508 (11.8)		589 (11.9)	517 (11.5)	131 (11.3)	

SD: standard deviation; [†]: P values from analysis of variance;[‡]: P values from chi-square test

Table3 Observational multivariable linear or logistic regression estimate of the cross-sectional association of alanine aminotransferase (ALT, U/L) with cardiovascular disease (CVD) risk factors, self-reported ischemic heart disease and type 2 diabetes in 19,925GBCS participants from 2003 to 2008

	Crude model			**Multivariate regression model**†		
	β	**95% CI**	**P value**	**β**	**95% CI**	**P value**
Systolic blood pressure, mmHg	0.12	0.1 to 0.15	<0.001	0.03	0.01 to 0.05	<0.001
Diastolic blood pressure, mmHg	0.09	0.08 to 0.11	<0.001	0.02	0.01 to 0.03	<0.001
High density lipoprotein-cholesterol, mmol/l	-0.002	-0.002 to -0.001	<0.001	0.0001	-0.0002 to 0.0005	0.50
Low density lipoprotein-cholesterol, mmol/l	0.001	0.00004 to 0.0016	0.04	0.001	0.0004 to 0.0017	<0.001
Triglycerides, mmol/l	0.013	0.011 to 0.014	<0.001	0.009	0.007 to 0.01	<0.001
Total cholesterol, mmol/l	0.004	0.003 to 0.005	<0.001	0.003	0.0023 to 0.0043	<0.001
Fasting glucose, mmol/l	0.010	0.01 to 0.02	<0.001	0.008	0.01 to 0.01	<0.001
Body mass index, Kg/m^2	0.04	0.04 to 0.04	<0.001	0.04	0.04 to 0.04	<0.001
Waist circumference, cm	0.12	0.11 to 0.12	<0.001	0.11	0.10 to 0.12	<0.001
Corrected QT interval, milliseconds	0.01	-0.01 to 0.03	0.27	0.02	-0.0003 to 0.04	0.054
Uncorrected QT interval, milliseconds	-0.02	-0.05 to 0	0.08	-0.02	-0.04 to 0.01	0.27
Heart rate, beats per minute	0.02	0.01 to 0.03	<0.001	0.02	0.01 to 0.03	<0.001
	Odds ratio	**95% CI**	**P value**	**Odds ratio**	**95% CI**	**P value**
Ischemic heart disease	1.005	1.00 to 1.01	0.05	1.0009	0.996 to 1.01	0.71
Type 2 diabetes	1.02	1.019 to 1.02	<0.001	1.012	1.009 to 1.01	<0.001

CI: confidence interval

†: Adjusted for age, sex, education, smoking status, alcohol use, physical activity and adiposity (BMI and waist circumference, except for the results for BMI and waist circumference)

Table 4 Mendelian randomization estimates, obtained from instrumental variable (IV) analysis using data from Guangzhou Biobank Cohort Study (GBCS) of the association of alanine aminotransferase (ALT) with cardiovascular disease (CVD) risk factors, Ischemic heart disease and type 2 diabetes.

	GBCS MR analysis (n=10,623)		
	β[†‡]	**95% CI**	**P value**
Systolic blood pressure, mmHg	0.19	-0.51 to 0.89	0.60
Diastolic blood pressure, mmHg	0.06	-0.29 to 0.42	0.74
High density lipoprotein-cholesterol, mmol/l	-0.0001	-0.01 to 0.01	0.98
Low density lipoprotein-cholesterol, mmol/l	0.01	-0.01 to 0.03	0.40
Triglycerides, mmol/l	-0.08	-0.13 to -0.03	0.004
Total cholesterol, mmol/l	-0.01	-0.05 to 0.02	0.44
Fasting glucose, mmol/l	0.03	-0.02 to 0.07	0.29
Body mass index, Kg/m^2	-0.05	-0.15 to 0.06	0.40
Waist circumference, cm	-0.12	-0.41 to 0.17	0.43
Corrected QT interval, milliseconds	0.44	-0.24 to 1.12	0.21
Uncorrected QT interval, milliseconds	0.31	-0.59 to 1.2	0.50
Heart rate, beats per minute	0.03	-0.3 to 0.36	0.88
	Odds ratio	**95% CI**	**P value**
Ischemic heart disease	0.92	0.87 to 0.97	0.004
Type 2 Diabetes	0.99	0.91 to 1.08	0.87

[†]: P-values from the Durbin-Wu-Hausman endogeneity test were non-significant (from 0.07 to 1.00) for the cardiovascular risk factors above except for triglycerides (P<0.001).

[‡]: Beta coefficients reflect differences in mean outcomes per 1 U/L difference of ALT